

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Boynton, John E.
Gillham, Nicholas W.
Randolph-Anderson, Barbara L.
Ishige, Fumiharu
Sato, Ryo

(ii) TITLE OF INVENTION: Methods of Conferring PPO-Inhibiting
Herbicide Resistance to Plants by Gene Manipulation

(iii) NUMBER OF SEQUENCES: 24

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
(B) STREET: P.O. Box 747
(C) CITY: Falls Church
(D) STATE: Virginia
(E) COUNTRY: USA
(F) ZIP: 22040-3487

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US new
(B) FILING DATE: 30-SEP-1996
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Murphy Jr., Gerald M.
(B) REGISTRATION NUMBER: 28,977
(C) REFERENCE/DOCKET NUMBER: 2185-156P

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 703-205-8000
(B) TELEFAX: 703-205-8050

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 47 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

65

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Chlamydomonas reinhardtii*
- (B) STRAIN: CC-407

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..47

(D) OTHER INFORMATION: /product= "porphyric herbicide resistance domain"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Ala Ala Glu Ala Leu Gly Ser Phe Asp Tyr Pro Pro Val Gly Ala Val
1 5 10 15

Thr Leu Ser Tyr Pro Leu Ser Ala Val Arg Glu Glu Arg Lys Ala Ser
20 25 30

Asp Gly Ser Val Pro Gly Phe Gly Gln Leu His Pro Arg Thr Gln
35 40 45

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 46 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Arabidopsis thaliana*
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..46

(D) OTHER INFORMATION: /product= "porphyric herbicide resistance domain"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Ala Ala Asn Ala Leu Ser Lys Leu Tyr Tyr Pro Pro Val Ala Ala Val
1 5 10 15

Ser Ile Ser Tyr Pro Lys Glu Ala Ile Arg Thr Glu Cys Leu Ile Asp
20 25 30

Gly Glu Leu Lys Gly Phe Gly Gln Leu His Pro Arg Thr Gln
35 40 45

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 46 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Zea mays*
- (B) STRAIN: B73 inbred

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..46

(D) OTHER INFORMATION: /product= "porphyric herbicide
resistance domain"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Ala Ala Asp Ala Leu Ser Arg Phe Tyr Tyr Pro Pro Val Ala Ala Val
1 5 10 15

Thr Val Ser Tyr Pro Lys Glu Ala Ile Arg Lys Glu Cys Leu Ile Asp
20 25 30

Gly Glu Leu Gln Gly Phe Gly Gln Leu His Pro Arg Ser Gln
35 40 45

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 141 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Chlamydomonas reinhardtii*
- (B) STRAIN: CC-407

(ix) FEATURE:

- (A) NAME/KEY: -
- (B) LOCATION: 1..141
- (D) OTHER INFORMATION: /note= "encodes porphyric herbicide resistance domain"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

| | |
|---|-----|
| GCCGCCGAGG CCCTGGGCTC CTTCGACTAC CCGCCGGTGG GCGCCGTGAC GCTGTCGTAC | 60 |
| CCGCTGAGCG CCGTGCAGGA GGAGCGCAAG GCCTCGGACG GGTCCGTGCC GGGCTTCGGT | 120 |
| CAGCTGCACC CGCGCACGCA G | 141 |

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 138 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Arabidopsis thaliana*
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: -
- (B) LOCATION: 1..138
- (D) OTHER INFORMATION: /note= "encodes porphyric herbicide resistance domain"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

| | |
|--|-----|
| GCTGCAAATG CACTCTAAA ACTATATTAC CCACCAGTTG CAGCAGTATC TATCTCGTAC | 60 |
| CCGAAAGAAG CAATCCGAAC AGAATGTTG ATAGATGGTG AACTAAAGGG TTTTGGCAA | 120 |
| TTGCATCCAC GCACGCAA | 138 |

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 138 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Zea mays*
- (B) STRAIN: B73 inbred

(ix) FEATURE:

- (A) NAME/KEY: -
- (B) LOCATION: 1..138
- (D) OTHER INFORMATION: /note= "encodes porphyric herbicide resistance domain"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

| | |
|---|-----|
| GCTGCAGATG CTCTATCAAG ATTCTATTAT CCACCGGTTG CTGCTGTAAC TGTTTCGTAT | 60 |
| CCAAAGGAAG CAATTAGAAA AGAATGCTTA ATTGATGGGG AACTCCAGGG CTTTGGCCAG | 120 |
| TTGCATCCAC GTAGTCAA | 138 |

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "oligonucleotide"

(iii) HYPOTHETICAL: NO

(ix) FEATURE:

- (A) NAME/KEY: -
- (B) LOCATION: 1..36
- (C) OTHER INFORMATION:/NOTE = "oligonucleotide primer for *Arabidopsis thaliana*"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

| | |
|---|----|
| CTATATTACC CACCAATGGC AGCAGTATCT ATCTCG | 36 |
|---|----|

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: "oligonucleotide"

(iii) HYPOTHETICAL: NO
(ix) FEATURE:
(A) NAME/KEY: -
(B) LOCATION: 1..38
(C) OTHER INFORMATION:/NOTE = "oligonucleotide primer for Zea mays"
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GATTCTATTA TCCACCGATG GCTGCTGTAA CTGTTTCG

38

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: "oligonucleotide"

(iii) HYPOTHETICAL: YES

(ix) FEATURE:

(A) NAME/KEY: -
(B) LOCATION: 1..26
(D) OTHER INFORMATION: /note= "oligonucleotide primer common to both of A. thaliana and Z. mays porphyric herbicide resistance domain of PPO."
/note= "N residues can be inosine
(I) in addition to G, A, T or C. K = G or T, Y = C or T, S = C or G,
W = A or T

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

KAYTAYCCNC CNATGGSNGC NGTNWS

26

(2) INFORMATION FOR SEQ ID NO:10

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2573 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: not relevant

(ii) MOLECULAR TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *Chlamydomonas reinhardtii*
(B) STRAIN: RS-3

(ix) FEATURE:
(A) NAME/KEY: -
(B) LOCATION: 1..2573

(C) OTHER INFORMATION:/note="encodes protoporphyrinogen oxidase"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

| | |
|---|------|
| CTCGAGAGCG TTGGAGGAAA TCCGTTGGC ACCTGTTCCG GCTTCTTGT GTGCACGGCC | 60 |
| ACGTCCCCCT TTCCCTGCTAC CCGCTCCCCC CCGGCTTTAC TGCCCTTCC ACTCCTCGGC | 120 |
| TCCATCCCGA TTCCATCCGC TCCTCCTCCC CCACCTAGAC TGTCTACCGT CTACCAGTT | 180 |
| CTTGGGCAAT CATTAAACGTA ACCCCGCGCTC CCTGCGCCTG CCCCTCCCTC CCTCTCCCCC | 240 |
| CCGCACAGCC CGCCGCCGCC GAGGCCCTGG GCTCCTTCGA CTACCCGCCG ATGGGCCCG | 300 |
| TGACGCTGTC GTACCCGCTG AGCGCCGTGC GGGAGGAGCG CAAGGCCCTCG GACGGGTCCG | 360 |
| TGCCGGGCTT CGGTCAGCTG CACCCGCGCA CGCAGGTGGG CAAGTGCACG CGTGTGCGG | 420 |
| GCGGTGTGTT GCGGAGGGGA GGGTGGTGGG GGTGTTGGGGT GGGGGTGGGG GGGATTGGGG | 480 |
| CGCTGGGTCG TATCCCGCGG TTGTATCCTC GCGCTCCCCCT CATCCATTCC CCCCTTCAAC | 540 |
| AACACACACG GGCGCACACG CACCCCTTTT GCGCTTACTT TGTCTGGTGC TCCTTAACAC | 600 |
| ACTCTCGCT TCATTTGGT GTCTTCTAAC ACACACACTT GTCCACACAC AGGGCATCAC | 660 |
| CACTCTGGGC ACCATCTACA GCTCCAGCCT GTTCCCCGGC CGCGCGCCCG AGGGCCACAT | 720 |
| GCTGCTGCTC AACTACATCG GCGGCACAC CAACCGCGGC ATCGTCAACC AGACCACCGA | 780 |
| GCAGCTGGTG GAGCAGGTGT GTGTGTGGG GGGTGGGGGG GGGGCAGTGG ATTTTTGGGC | 840 |
| TGAGCCCCCT GAGCAAAGCG ATCCAGGGGG GGCAGAGCCC CCCAGGATTG CCCCTGTCCG | 900 |
| TGCGTGCCTG TGTGCCTGTG TCGACAAAAA GTACCGTACT GGCACAAACC GCGAGTGCCA | 960 |
| CGTATTATTA ATTGCAATTAA CCTATTGTAG AAAAATAGAC GGCAGGGAAA ACTCGGCCGG | 1020 |
| AGCGAGAACG GACCTCGTGA GTCCATGGAC ATCTTGACTT TCTTCAGTTC GCGAGTATAG | 1080 |
| CTCTCGGCCCT CTAATATCT TACATCCATG TATCAAAACA TGTCGACGAC AAGCGTCTTG | 1140 |
| GGGCAAGAAT GTCGAAATTG TTTGCAACAG CCAAACCATG CGTCCCCGAG CCTTACATGT | 1200 |
| GTCGCGGCCCG GGGATCCCGC GCCCGAGCCC GGCTAGCCCT TTGCGGTGCT TGAGTGGGAT | 1260 |
| GTGGGTGAGG TGCATTTGGG ATATCATGGA CCGTGAAGTG GCGTGGGTAA GGTGGCGTGG | 1320 |
| CGTGGCGGGG ACAGGGCATG TCGGTGCCTC GGCACAGCGT TGGCCTAGTG GCCAGTCCCG | 1380 |
| CTGGATGGGC TTGCAAGGGT GCTGTTCATG TCGCCGGTGC CCATCGTCAC ATCCCCTTGC | 1440 |
| GCTACATGGG GCTCAGCCCA TTTTCCAGCT GTACAAAGCT GACACCCCTT GTTGTGTGGC | 1500 |
| GTCTTGGACC CGTGTGCTT CGGAGCTGGC CAGAACCCCCC TGTGGGCACA CACACGACA | 1560 |

| | | | | | | | |
|------------|------------|------------|------------|-------------|------------|------------|------|
| CACACACACA | CACACACACA | CACACACACA | CACACACACA | CACACACACA | CACACACACA | CACACACACA | 1620 |
| CACACACACA | CACACACACA | CACACACACA | CACACACACA | CACACACACA | CACATTTCG | TCCTGCAGCC | 1680 |
| CCGAACCCCG | CCGCCCGTTC | CACGTCTTCC | ACCTGCCGCA | CCCCCCCCCC | TGCCGCACGC | | 1740 |
| CTGCTCTCAC | CGCCTCTCCC | CCCACCCAT | CTCCCTGCAG | GTGGACAAGG | ACCTGCGCAA | | 1800 |
| CATGGTCATC | AAGCCCGACG | CGCCCAAGCC | CCGTGTGGTG | GGCGTGCAGCG | TGTGGCCGCG | | 1860 |
| CGCCATCCCG | CAGGTGTGAG | GGCGCAGCAG | CCGGAGGGAT | GGGCTAGATC | CTAGTTCTC | | 1920 |
| AAAGAGCTCT | ACAGCCCTAT | AACCTCGACC | TGCGACCTTC | GACCTGATAA | CCTGGCTGCC | | 1980 |
| CCCTCCCAAC | CTAGCCACCT | CTCCCCGGAT | TTGGGTTCAC | TCGGTTGACT | TGCTTTGGG | | 2040 |
| TTCTGGAATC | AACTTCACCT | GTTGTATACT | TTGCTGCACT | TCTCTGTACC | ACTCTTGCA | | 2100 |
| TTAGGTTCGG | TTTAGTTGG | GCTGCATGTG | TAACCCCTCC | TCCCCGCCCT | GCCACCTGCA | | 2160 |
| GTTCAACCTG | GGCCACCTGG | AGCAGCTGGA | CAAGGCGCGC | AAGGCGCTGG | ACGCGGCGGG | | 2220 |
| GCTGCAGGGC | GTGCACCTGG | GGGGCAACTA | CGTCAGCGGT | GAGCGCGTGG | GCAGCAGCAG | | 2280 |
| CAGCAGGAAG | AGGGGAGGGG | AGGGGAGGGG | AGGGTACAAG | GAGGAGGTTG | AGCAGGAGGT | | 2340 |
| GGTGCTAAGG | CGCAAAGCAA | GGCGGTGTTG | TATCCTCATT | GAUTGAAACC | GGGAAACCCA | | 2400 |
| GCATGAACAA | GAGGTCAAGG | GAUTGCAAGG | AGCGGAGGCT | ACATGTATGA | CTACCCCCGA | | 2460 |
| CGCGGGCGAT | GATTCTTGA | CTATTGGAC | CTATTCGTT | GGGCTCGGGC | ACATGACCCC | | 2520 |
| CCTGGCCCT | TCGCTGTATG | GTGCCAGCC | GCCCAGCCGC | CCCCGGCCCA | CAC | | 2573 |

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1704 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Arabidopsis thaliana*
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 16..1629
- (D) OTHER INFORMATION: /product= "protoporphyrinogen oxidase"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

| | |
|--|-----|
| TTCTCTGCGA TTTCC ATG GAG TTA TCT CTT CTC CGT CCG ACG ACT CAA TCG | 51 |
| Met Glu Leu Ser Leu Leu Arg Pro Thr Thr Gln Ser | |
| 1 5 10 | |
| CTT CTT CCG TCG TTT TCG AAG CCC AAT CTC CGA TTA AAT GTT TAT AAG | 99 |
| Leu Leu Pro Ser Phe Ser Lys Pro Asn Leu Arg Leu Asn Val Tyr Lys | |
| 15 20 25 | |
| CCT CTT AGA CTC CGT TGT TCA GTG GCC GGT GGA CCA ACC GTC GGA TCT | 147 |
| Pro Leu Arg Leu Arg Cys Ser Val Ala Gly Gly Pro Thr Val Gly Ser | |
| 30 35 40 | |
| TCA AAA ATC GAA GGC GGA GGA GGC ACC ACC ATC ACG ACG GAT TGT GTG | 195 |
| Ser Lys Ile Glu Gly Gly Thr Thr Ile Thr Thr Asp Cys Val | |
| 45 50 55 60 | |
| ATT GTC GGC GGA GGT ATT AGT GGT CTT TGC ATC GCT CAG GCG CTT GCT | 243 |
| Ile Val Gly Gly Ile Ser Gly Leu Cys Ile Ala Gln Ala Leu Ala | |
| 65 70 75 | |
| ACT AAG CAT CCT GAT GCT GCT CCG AAT TTA ATT GTG ACC GAG GCT AAG | 291 |
| Thr Lys His Pro Asp Ala Ala Pro Asn Leu Ile Val Thr Glu Ala Lys | |
| 80 85 90 | |
| GAT CGT GTT GGA GGC AAC ATT ATC ACT CGT GAA GAG AAT GGT TTT CTC | 339 |
| Asp Arg Val Gly Gly Asn Ile Ile Thr Arg Glu Glu Asn Gly Phe Leu | |
| 95 100 105 | |
| TGG GAA GAA GGT CCC AAT AGT TTT CAA CCG TCT GAT CCT ATG CTC ACT | 387 |
| Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp Pro Met Leu Thr | |
| 110 115 120 | |
| ATG GTG GTA GAT AGT GGT TTG AAG GAT GAT TTG GTG TTG GGA GAT CCT | 435 |
| Met Val Val Asp Ser Gly Leu Lys Asp Asp Leu Val Leu Gly Asp Pro | |
| 125 130 135 140 | |
| ACT GCG CCA AGG TTT GTG TTG TGG AAT GGG AAA TTG AGG CCG GTT CCA | 483 |
| Thr Ala Pro Arg Phe Val Leu Trp Asn Gly Lys Leu Arg Pro Val Pro | |
| 145 150 155 | |
| TCG AAG CTA ACA GAC TTA CCG TTC TTT GAT TTG ATG AGT ATT GGT GGG | 531 |
| Ser Lys Leu Thr Asp Leu Pro Phe Phe Asp Leu Met Ser Ile Gly Gly | |
| 160 165 170 | |
| AAG ATT AGA GCT GGT TTT GGT GCA CTT GGC ATT CGA CCG TCA CCT CCA | 579 |
| Lys Ile Arg Ala Gly Phe Gly Ala Leu Gly Ile Arg Pro Ser Pro Pro | |
| 175 180 185 | |
| GGT CGT GAA GAA TCT GTG GAG GAG TTT GTA CGG CGT AAC CTC GGT GAT | 627 |
| Gly Arg Glu Glu Ser Val Glu Glu Phe Val Arg Arg Asn Leu Gly Asp | |
| 190 195 200 | |
| GAG GTT TTT GAG CGC CTG ATT GAA CCG TTT TGT TCA GGT GTT TAT GCT | 675 |

| | | | |
|---|-----|-----|------|
| Glu Val Phe Glu Arg Leu Ile Glu Pro Phe Cys Ser Gly Val Tyr Ala | | | |
| 205 | 210 | 215 | 220 |
| GGT GAT CCT TCA AAA CTG AGC ATG AAA GCA GCG TTT GGG AAG GTT TGG | | | 723 |
| Gly Asp Pro Ser Lys Leu Ser Met Lys Ala Ala Phe Gly Lys Val Trp | 225 | 230 | 235 |
| AAA CTA GAG CAA AAT GGT GGA AGC ATA ATA GGT GGT ACT TTT AAG GCA | | | 771 |
| Lys Leu Glu Gln Asn Gly Gly Ser Ile Ile Gly Gly Thr Phe Lys Ala | 240 | 245 | 250 |
| ATT CAG GAG AGG AAA AAC GCT CCC AAG GCA GAA CGA GAC CCG CGC CTG | | | 819 |
| Ile Gln Glu Arg Lys Asn Ala Pro Lys Ala Glu Arg Asp Pro Arg Leu | 255 | 260 | 265 |
| CCA AAA CCA CAG GGC CAA ACA GTT GGT TCT TTC AGG AAG GGA CTT CGA | | | 867 |
| Pro Lys Pro Gln Gly Gln Thr Val Gly Ser Phe Arg Lys Gly Leu Arg | 270 | 275 | 280 |
| ATG TTG CCA GAA GCA ATA TCT GCA AGA TTA GGT AGC AAA GTT AAG TTG | | | 915 |
| Met Leu Pro Glu Ala Ile Ser Ala Arg Leu Gly Ser Lys Val Lys Leu | 285 | 290 | 295 |
| 300 | | | |
| TCT TGG AAG CTC TCA GGT ATC ACT AAG CTG GAG AGC GGA GGA TAC AAC | | | 963 |
| Ser Trp Lys Leu Ser Gly Ile Thr Lys Leu Glu Ser Gly Gly Tyr Asn | 305 | 310 | 315 |
| TTA ACA TAT GAG ACT CCA GAT GGT TTA GTT TCC GTG CAG AGC AAA AGT | | | 1011 |
| Leu Thr Tyr Glu Thr Pro Asp Gly Leu Val Ser Val Gln Ser Lys Ser | 320 | 325 | 330 |
| 330 | | | |
| GTT GTA ATG ACG GTG CCA TCT CAT GTT GCA AGT GGT CTC TTG CGC CCT | | | 1059 |
| Val Val Met Thr Val Pro Ser His Val Ala Ser Gly Leu Leu Arg Pro | 335 | 340 | 345 |
| 345 | | | |
| CTT TCT GAA TCT GCT GCA AAT GCA CTC TCA AAA CTA TAT TAC CCA CCA | | | 1107 |
| Leu Ser Glu Ser Ala Ala Asn Ala Leu Ser Lys Leu Tyr Tyr Pro Pro | 350 | 355 | 360 |
| 360 | | | |
| GTT GCA GCA GTA TCT ATC TCG TAC CCG AAA GAA GCA ATC CGA ACA GAA | | | 1155 |
| Val Ala Ala Val Ser Ile Ser Tyr Pro Lys Glu Ala Ile Arg Thr Glu | 365 | 370 | 375 |
| 380 | | | |
| TGT TTG ATA GAT GGT GAA CTA AAG GGT TTT GGG CAA TTG CAT CCA CGC | | | 1203 |
| Cys Leu Ile Asp Gly Glu Leu Lys Gly Phe Gly Gln Leu His Pro Arg | 385 | 390 | 395 |
| 395 | | | |
| ACG CAA GGA GTT GAA ACA TTA GGA ACT ATC TAC AGC TCC TCA CTC TTT | | | 1251 |
| Thr Gln Gly Val Glu Thr Leu Gly Thr Ile Tyr Ser Ser Ser Leu Phe | 400 | 405 | 410 |
| 410 | | | |
| CCA AAT CGC GCA CCG CCC GGA AGA ATT TTG CTG TTG AAC TAC ATT GGC | | | 1299 |
| Pro Asn Arg Ala Pro Pro Gly Arg Ile Leu Leu Leu Asn Tyr Ile Gly | 415 | 420 | 425 |
| 425 | | | |
| GGG TCT ACA AAC ACC GGA ATT CTG TCC AAG TCT GAA GGT GAG TTA GTG | | | 1347 |

| | | | |
|---|-----|-----|------|
| Gly Ser Thr Asn Thr Gly Ile Leu Ser Lys Ser Glu Gly Glu Leu Val | | | |
| 430 | 435 | 440 | |
| GAA GCA GTT GAC AGA GAT TTG AGG AAA ATG CTA ATT AAG CCT AAT TCG | | | 1395 |
| Glu Ala Val Asp Arg Asp Leu Arg Lys Met Leu Ile Lys Pro Asn Ser | | | |
| 445 | 450 | 455 | 460 |
| ACC GAT CCA CTT AAA TTA GGA GTT AGG GTA TGG CCT CAA GCC ATT CCT | | | |
| Thr Asp Pro Leu Lys Leu Gly Val Arg Val Trp Pro Gln Ala Ile Pro | | | |
| 465 | 470 | | 475 |
| CAG TTT CTA GTT GGT CAC TTT GAT ATC CTT GAC ACG GCT AAA TCA TCT | | | 1443 |
| Gln Phe Leu Val Gly His Phe Asp Ile Leu Asp Thr Ala Lys Ser Ser | | | |
| 480 | 485 | | 490 |
| CTA ACG TCT TCG GGC TAC GAA GGG CTA TTT TTG GGT GGC AAT TAC GTC | | | 1491 |
| Leu Thr Ser Ser Gly Tyr Glu Gly Leu Phe Leu Gly Gly Asn Tyr Val | | | |
| 495 | 500 | | 505 |
| GCT GGT GTA GCC TTA GGC CGG TGT GTA GAA GGC GCA TAT GAA ACC GCG | | | 1539 |
| Ala Gly Val Ala Leu Gly Arg Cys Val Glu Gly Ala Tyr Glu Thr Ala | | | |
| 510 | 515 | | 520 |
| ATT GAG GTC AAC AAC TTC ATG TCA CGG TAC GCT TAC AAG TAA | | | 1587 |
| Ile Glu Val Asn Asn Phe Met Ser Arg Tyr Ala Tyr Lys * | | | |
| 525 | 530 | | 535 |
| ATGTAAAACA TTAAATCTCC CAGCTTGCCT GAGTTTATT AAATATTTG AGATATCCAA | | | 1629 |
| AAAAAAAAAAAA AAAAAA | | | 1689 |
| | | | 1704 |

(2) INFORMATION FOR SEQ ID NO:12

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 537 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: protein

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Arabidopsis thaliana*
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..537
- (C) OTHER INFORMATION: /product="protoporphyrinogen oxidase"

(xi) SEQUENCE DESCRIPTION: SEQ. ID. NO:12:

| | | | |
|---|---|----|----|
| Met Glu Leu Ser Leu Leu Arg Pro Thr Thr Gln Ser Leu Leu Pro Ser | | | |
| 1 | 5 | 10 | 15 |

75

Phe Ser Lys Pro Asn Leu Arg Leu Asn Val Tyr Lys Pro Leu Arg Leu
 20 25 30

Arg Cys Ser Val Ala Gly Gly Pro Thr Val Gly Ser Ser Lys Ile Glu
 35 40 45

Gly Gly Gly Thr Thr Ile Thr Thr Asp Cys Val Ile Val Gly Gly
 50 55 60

Gly Ile Ser Gly Leu Cys Ile Ala Gln Ala Leu Ala Thr Lys His Pro
 65 70 75 80

Asp Ala Ala Pro Asn Leu Ile Val Thr Glu Ala Lys Asp Arg Val Gly
 85 90 95

Gly Asn Ile Ile Thr Arg Glu Glu Asn Gly Phe Leu Trp Glu Glu Gly
 100 105 110

Pro Asn Ser Phe Gln Pro Ser Asp Pro Met Leu Thr Met Val Val Asp
 115 120 125

Ser Gly Leu Lys Asp Asp Leu Val Leu Gly Asp Pro Thr Ala Pro Arg
 130 135 140

Phe Val Leu Trp Asn Gly Lys Leu Arg Pro Val Pro Ser Lys Leu Thr
 145 150 155 160

Asp Leu Pro Phe Phe Asp Leu Met Ser Ile Gly Gly Lys Ile Arg Ala
 165 170 175

Gly Phe Gly Ala Leu Gly Ile Arg Pro Ser Pro Pro Gly Arg Glu Glu
 180 185 190

Ser Val Glu Glu Phe Val Arg Arg Asn Leu Gly Asp Glu Val Phe Glu
 195 200 205

Arg Leu Ile Glu Pro Phe Cys Ser Gly Val Tyr Ala Gly Asp Pro Ser
 210 215 220

Lys Leu Ser Met Lys Ala Ala Phe Gly Lys Val Trp Lys Leu Glu Gln
 225 230 235 240

Asn Gly Gly Ser Ile Ile Gly Gly Thr Phe Lys Ala Ile Gln Glu Arg
 245 250 255

Lys Asn Ala Pro Lys Ala Glu Arg Asp Pro Arg Leu Pro Lys Pro Gln
 260 265 270

Gly Gln Thr Val Gly Ser Phe Arg Lys Gly Leu Arg Met Leu Pro Glu
 275 280 285

Ala Ile Ser Ala Arg Leu Gly Ser Lys Val Lys Leu Ser Trp Lys Leu
 290 295 300

Ser Gly Ile Thr Lys Leu Glu Ser Gly Gly Tyr Asn Leu Thr Tyr Glu

76

| | | | |
|---|-----|-----|-----|
| 305 | 310 | 315 | 320 |
| Thr Pro Asp Gly Leu Val Ser Val Gln Ser Lys Ser Val Val Met Thr | | | |
| 325 | 330 | 335 | |
| Val Pro Ser His Val Ala Ser Gly Leu Leu Arg Pro Leu Ser Glu Ser | | | |
| 340 | 345 | 350 | |
| Ala Ala Asn Ala Leu Ser Lys Leu Tyr Tyr Pro Pro Val Ala Ala Val | | | |
| 355 | 360 | 365 | |
| Ser Ile Ser Tyr Pro Lys Glu Ala Ile Arg Thr Glu Cys Leu Ile Asp | | | |
| 370 | 375 | 380 | |
| Gly Glu Leu Lys Gly Phe Gly Gln Leu His Pro Arg Thr Gln Gly Val | | | |
| 385 | 390 | 395 | 400 |
| Glu Thr Leu Gly Thr Ile Tyr Ser Ser Ser Leu Phe Pro Asn Arg Ala | | | |
| 405 | 410 | 415 | |
| Pro Pro Gly Arg Ile Leu Leu Leu Asn Tyr Ile Gly Gly Ser Thr Asn | | | |
| 420 | 425 | 430 | |
| Thr Gly Ile Leu Ser Lys Ser Glu Gly Glu Leu Val Glu Ala Val Asp | | | |
| 435 | 440 | 445 | |
| Arg Asp Leu Arg Lys Met Leu Ile Lys Pro Asn Ser Thr Asp Pro Leu | | | |
| 450 | 455 | 460 | |
| Lys Leu Gly Val Arg Val Trp Pro Gln Ala Ile Pro Gln Phe Leu Val | | | |
| 465 | 470 | 475 | 480 |
| Gly His Phe Asp Ile Leu Asp Thr Ala Lys Ser Ser Leu Thr Ser Ser | | | |
| 485 | 490 | 495 | |
| Gly Tyr Glu Gly Leu Phe Leu Gly Gly Asn Tyr Val Ala Gly Val Ala | | | |
| 500 | 505 | 510 | |
| Leu Gly Arg Cys Val Glu Gly Ala Tyr Glu Thr Ala Ile Glu Val Asn | | | |
| 515 | 520 | 525 | |
| Asn Phe Met Ser Arg Tyr Ala Tyr Lys * | | | |
| 530 | 535 | | |

(2) INFORMATION FOR SEQ ID NO:13

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1698 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: not relevant

(ii) MOLECULAR TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Zea mays*
 (B) STRAIN: B73 inbred

(ix) FEATURE:

(A) NAME/KEY: CDS
 (B) LOCATION: 2..1453
 (C) OTHER INFORMATION: /product="protoporphyrinogen oxidase"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

| | |
|---|-----|
| G AAT TCG GCG GAC TGC GTC GTG GTG GGC GGA GGC ATC AGT GGC CTC | 46 |
| Asn Ser Ala Asp Cys Val Val Val Gly Gly Gly Ile Ser Gly Leu | |
| 1 5 10 15 | |
| TGC ACC GCG CAG GCG CTG GCC ACG CGG CAC GGC GTC GGG GAC GTG CTT | 94 |
| Cys Thr Ala Gln Ala Leu Ala Thr Arg His Gly Val Gly Asp Val Leu | |
| 20 25 30 | |
| GTC ACG GAG GCC CGC GCC CGC CCC GGC GGC AAC ATT ACC ACC GTC GAG | 142 |
| Val Thr Glu Ala Arg Ala Arg Pro Gly Gly Asn Ile Thr Thr Val Glu | |
| 35 40 45 | |
| CGC CCC GAG GAA GGG TAC CTC TGG GAG GAG GGT CCC AAC AGC TTC CAG | 190 |
| Arg Pro Glu Glu Gly Tyr Leu Trp Glu Glu Gly Pro Asn Ser Phe Gln | |
| 50 55 60 | |
| — GGC TCC GAC CCC GTT CTC ACC ATG GCC GTG GAC AGC GGA CTG AAG GAT | 238 |
| Pro Ser Asp Pro Val Leu Thr Met Ala Val Asp Ser Gly Leu Lys Asp | |
| 65 70 75 | |
| GAC TTG GTT TTT GGG GAC CCA AAC GCG CCG CGT TTC GTG CTG TGG GAG | 286 |
| Asp Leu Val Phe Gly Asp Pro Asn Ala Pro Arg Phe Val Leu Trp Glu | |
| 80 85 90 95 | |
| GGG AAG CTG AGG CCC GTG CCA TCC AAG CCC GCC GAC CTC CCG TTC TTC | 334 |
| Gly Lys Leu Arg Pro Val Pro Ser Lys Pro Ala Asp Leu Pro Phe Phe | |
| 100 105 110 | |
| GAT CTC ATG AGC ATC CCA GGG AAG CTC AGG GCC GGT CTA GGC GCG CTT | 382 |
| Asp Leu Met Ser Ile Pro Gly Lys Leu Arg Ala Gly Leu Gly Ala Leu | |
| 115 120 125 | |
| GGC ATC CGC CCG CCT CCA GGC CGC GAA GAG TCA GTG GAG GAG TTC | 430 |
| Gly Ile Arg Pro Pro Pro Gly Arg Glu Glu Ser Val Glu Glu Phe | |
| 130 135 140 | |
| GTG CGC CGC AAC CTC GGT GCT GAG GTC TTT GAG CGC CTC ATT GAG CCT | 478 |
| Val Arg Arg Asn Leu Gly Ala Glu Val Phe Glu Arg Leu Ile Glu Pro | |
| 145 150 155 | |
| TTC TGC TCA GGT GTC TAT GCT GGT GAT CCT TCT AAG CTC AGC ATG AAG | 526 |
| Phe Cys Ser Gly Val Tyr Ala Gly Asp Pro Ser Lys Leu Ser Met Lys | |
| 160 165 170 175 | |
| GCT GCA TTT GGG AAG GTT TGG CGG TTG GAA GAA ACT GGA GGT AGT ATT | 574 |

| | | | | |
|---|-----|-----|-----|------|
| Ala Ala Phe Gly Lys Val Trp Arg Leu Glu Glu Thr Gly Gly Ser Ile | 180 | 185 | 190 | |
| ATT GGT GGA ACC ATC AAG ACA ATT CAG GAG AGG AGC AAG AAT CCA AAA | | | | 622 |
| Ile Gly Gly Thr Ile Lys Thr Ile Gln Glu Arg Ser Lys Asn Pro Lys | 195 | 200 | 205 | |
| CCA CCG AGG GAT GCC CGC CTT CCG AAG CCA AAA GGG CAG ACA GTT GCA | | | | 670 |
| Pro Pro Arg Asp Ala Arg Leu Pro Lys Pro Lys Gly Gln Thr Val Ala | 210 | 215 | 220 | |
| TCT TTC AGG AAG GGT CTT GCC ATG CTT CCA AAT GCC ATT ACA TCC AGC | | | | 718 |
| Ser Phe Arg Lys Gly Leu Ala Met Leu Pro Asn Ala Ile Thr Ser Ser | 225 | 230 | 235 | |
| TTG GGT AGT AAA GTC AAA CTA TCA TGG AAA CTC ACG AGC ATT ACA AAA | | | | 766 |
| Leu Gly Ser Lys Val Lys Leu Ser Trp Lys Leu Thr Ser Ile Thr Lys | 240 | 245 | 250 | 255 |
| TCA GAT GAC AAG GGA TAT GTT TTG GAG TAT GAA ACG CCA GAA GGG GTT | | | | 814 |
| Ser Asp Asp Lys Gly Tyr Val Leu Glu Tyr Glu Thr Pro Glu Gly Val | 260 | 265 | 270 | |
| GTT TCG GTG CAG GCT AAA AGT GTT ATC ATG ACT ATT CCA TCA TAT GTT | | | | 862 |
| Val Ser Val Gln Ala Lys Ser Val Ile Met Thr Ile Pro Ser Tyr Val | 275 | 280 | 285 | |
| GCT AGC AAC ATT TTG CGT CCA CTT TCA AGC GAT GCT GCA GAT GCT CTA | | | | 910 |
| Ala Ser Asn Ile Leu Arg Pro Leu Ser Ser Asp Ala Ala Asp Ala Leu | 290 | 295 | 300 | |
| TCA AGA TTC TAT TAT CCA CCG GTT GCT GCT GTA ACT GTT TCG TAT CCA | | | | 958 |
| Ser Arg Phe Tyr Tyr Pro Pro Val Ala Ala Val Thr Val Ser Tyr Pro | 305 | 310 | 315 | |
| AAG GAA GCA ATT AGA AAA GAA TGC TTA ATT GAT GGG GAA CTC CAG GGC | | | | 1006 |
| Lys Glu Ala Ile Arg Lys Glu Cys Leu Ile Asp Gly Glu Leu Gln Gly | 320 | 325 | 330 | 335 |
| TTT GGC CAG TTG CAT CCA CGT AGT CAA GGA GTT GAG ACA TTA GGA ACA | | | | 1054 |
| Phe Gly Gln Leu His Pro Arg Ser Gln Gly Val Glu Thr Leu Gly Thr | 340 | 345 | 350 | |
| ATA TAC AGT TCC TCA CTC TTT CCA AAT CGT GCT CCT GAC GGT AGG GTG | | | | 1102 |
| Ile Tyr Ser Ser Leu Phe Pro Asn Arg Ala Pro Asp Gly Arg Val | 355 | 360 | 365 | |
| TTA CTT CTA AAC TAC ATA GGA GGT GCT ACA AAC ACA GGA ATT GTT TCC | | | | 1150 |
| Leu Leu Leu Asn Tyr Ile Gly Gly Ala Thr Asn Thr Gly Ile Val Ser | 370 | 375 | 380 | |
| AAG ACT GAA AGT GAG CTG GTC GAA GCA GTT GAC CGT GAC CTC CGA AAA | | | | 1198 |
| Lys Thr Glu Ser Glu Leu Val Glu Ala Val Asp Arg Asp Leu Arg Lys | 385 | 390 | 395 | |
| ATG CTT ATA AAT TCT ACA GCA GTG GAC CCT TTA GTC CTT GGT GTT CGA | | | | 1246 |

79

| | | | | |
|--|-----|-----|-----|------|
| Met Leu Ile Asn Ser Thr Ala Val Asp Pro Leu Val Leu Gly Val Arg | | | | |
| 400 | 405 | 410 | 415 | |
| GTT TGG CCA CAA GCC ATA CCT CAG TTC CTG GTA GGA CAT CTT GAT CTT | | | | 1294 |
| Val Trp Pro Gln Ala Ile Pro Gln Phe Leu Val Gly His Leu Asp Leu | 420 | 425 | 430 | |
| CTG GAA GCC GCA AAA GCT GCC CTG GAC CGA GGT GGC TAC GAT GGG CTG | | | | 1342 |
| Leu Glu Ala Ala Lys Ala Ala Leu Asp Arg Gly Gly Tyr Asp Gly Leu | 435 | 440 | 445 | |
| TTC CTA GGA GGG AAC TAT GTT GCA GGA GTT GCC CTG GGC AGA TGC GTT | | | | 1390 |
| Phe Leu Gly Gly Asn Tyr Val Ala Gly Val Ala Leu Gly Arg Cys Val | 450 | 455 | 460 | |
| GAG GGC GCG TAT GAA AGT GCC TCG CAA ATA TCT GAC TTC TTG ACC AAG | | | | 1438 |
| Glu Gly Ala Tyr Glu Ser Ala Ser Gln Ile Ser Asp Phe Leu Thr Lys | 465 | 470 | 475 | |
| TAT GCC TAC AAG TGA TGAAAGAAGT GGAGCGCTAC TTGCCAATCG TTTATGTTGC | | | | 1493 |
| Tyr Ala Tyr Lys * | 480 | | | |
| ATAGATGAGG TGCCTCCGGG GAAAAAAAAG CTTGAATAGT ATTTTTTATT CTTATTTGT | | | | 1553 |
| AAATTGCATT TCTGTTCTTT TTTCTATCAG TAATTAGTTA TATTTAGTT CTGTAGGAGA | | | | 1613 |
| TTGTTCTGTT CACTGCCCTT CAAAAGAAAT TTTATTTTC ATTCTTTAT GAGAGCTGTG | | | | 1673 |
| CTACTTAAAA AAAAAAAA AAAAAA | | | | 1698 |

(2) INFORMATION FOR SEQ ID NO:14

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 483 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: protein

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Zea mays*
- (B) STRAIN: B73 inbred

(ix) FEATURE:

- (A) NAME/KEY: peptide
- (B) LOCATION: 1..483
- (C) OTHER INFORMATION: /note="protoporphyrinogen oxidase"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

| | | | |
|---|---|----|----|
| Asn Ser Ala Asp Cys Val Val Gly Gly Gly Ile Ser Gly Leu Cys | | | |
| 1 | 5 | 10 | 15 |

80

Thr Ala Gln Ala Leu Ala Thr Arg His Gly Val Gly Asp Val Leu Val
 20 25 30

Thr Glu Ala Arg Ala Arg Pro Gly Gly Asn Ile Thr Thr Val Glu Arg
 35 40 45

Pro Glu Glu Gly Tyr Leu Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro
 50 55 60

Ser Asp Pro Val Leu Thr Met Ala Val Asp Ser Gly Leu Lys Asp Asp
 65 70 75 80

Leu Val Phe Gly Asp Pro Asn Ala Pro Arg Phe Val Leu Trp Glu Gly
 85 90 95

Lys Leu Arg Pro Val Pro Ser Lys Pro Ala Asp Leu Pro Phe Phe Asp
 100 105 110

Leu Met Ser Ile Pro Gly Lys Leu Arg Ala Gly Leu Gly Ala Leu Gly
 115 120 125

Ile Arg Pro Pro Pro Pro Gly Arg Glu Glu Ser Val Glu Glu Phe Val
 130 135 140

Arg Arg Asn Leu Gly Ala Glu Val Phe Glu Arg Leu Ile Glu Pro Phe
 145 150 155 160

Cys Ser Gly Val Tyr Ala Gly Asp Pro Ser Lys Leu Ser Met Lys Ala
 165 170 175

Ala Phe Gly Lys Val Trp Arg Leu Glu Glu Thr Gly Gly Ser Ile Ile
 180 185 190

Gly Gly Thr Ile Lys Thr Ile Gln Glu Arg Ser Lys Asn Pro Lys Pro
 195 200 205

Pro Arg Asp Ala Arg Leu Pro Lys Pro Lys Gly Gln Thr Val Ala Ser
 210 215 220

Phe Arg Lys Gly Leu Ala Met Leu Pro Asn Ala Ile Thr Ser Ser Leu
 225 230 235 240

Gly Ser Lys Val Lys Leu Ser Trp Lys Leu Thr Ser Ile Thr Lys Ser
 245 250 255

Asp Asp Lys Gly Tyr Val Leu Glu Tyr Glu Thr Pro Glu Gly Val Val
 260 265 270

Ser Val Gln Ala Lys Ser Val Ile Met Thr Ile Pro Ser Tyr Val Ala
 275 280 285

Ser Asn Ile Leu Arg Pro Leu Ser Ser Asp Ala Ala Asp Ala Leu Ser
 290 295 300

Arg Phe Tyr Tyr Pro Pro Val Ala Ala Val Thr Val Ser Tyr Pro Lys
 305 310 315 320

81

Glu Ala Ile Arg Lys Glu Cys Leu Ile Asp Gly Glu Leu Gln Gly Phe
 325 330 335

Gly Gln Leu His Pro Arg Ser Gln Gly Val Glu Thr Leu Gly Thr Ile
 340 345 350

Tyr Ser Ser Ser Leu Phe Pro Asn Arg Ala Pro Asp Gly Arg Val Leu
 355 360 365

Leu Leu Asn Tyr Ile Gly Gly Ala Thr Asn Thr Gly Ile Val Ser Lys
 370 375 380

Thr Glu Ser Glu Leu Val Glu Ala Val Asp Arg Asp Leu Arg Lys Met
 385 390 395 400

Leu Ile Asn Ser Thr Ala Val Asp Pro Leu Val Leu Gly Val Arg Val
 405 410 415

Trp Pro Gln Ala Ile Pro Gln Phe Leu Val Gly His Leu Asp Leu Leu
 420 425 430

Glu Ala Ala Lys Ala Ala Leu Asp Arg Gly Gly Tyr Asp Gly Leu Phe
 435 440 445

Leu Gly Gly Asn Tyr Val Ala Gly Val Ala Leu Gly Arg Cys Val Glu
 450 455 460

Gly Ala Tyr Glu Ser Ala Ser Gln Ile Ser Asp Phe Leu Thr Lys Tyr
 465 470 475 480

Ala Tyr Lys

(2) INFORMATION FOR SEQ ID NO:15

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: oligonucleotide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

- (A) NAME/KEY: -
- (B) LOCATION: 1..18
- (C) OTHER INFORMATION: /note="oligonucleotide primer 1A for *Chlamydomonas reinhardtii*"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15

CCGTCTACCA GTTTCTTG

(2) INFORMATION FOR SEQ ID NO:16

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: oligonucleotide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(ix) FEATURE:

- (A) NAME/KEY: -
- (B) LOCATION: 1..19
- (C) OTHER INFORMATION: /note="oligonucleotide primer 1B for *Chlamydomonas reinhardtii*"

(xi) SEQUENCE DESCRIPTION:SEQ ID NO:16

ATACAACCGC GGGATACGA

(2) INFORMATION FOR SEQ ID NO:17

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: oligonucleotide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

- (A) NAME/KEY: -
- (B) LOCATION: 1..17
- (C) OTHER INFORMATION: /note="oligonucleotide primer 2A for *Chlamydomonas reinhardtii*"

(xi) SEQUENCE DESCRIPTION:SEQ ID NO:17

ACTTTGTCTG GTGCTCC

(2) INFORMATION FOR SEQ ID NO:18

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (ii) MOLECULAR TYPE: oligonucleotide
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: YES
- (ix) FEATURE:
 - (A) NAME/KEY: -
 - (B) LOCATION: 1..17
 - (C) OTHER INFORMATION: /note="oligonucleotide primer 2B for *Chlamydomonas reinhardtii*"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18

TGGATCGCTT TGCTCAG

(2) INFORMATION FOR SEQ ID NO:19

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 3381 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: not relevant
 - (D) TOPOLOGY: not relevant
- (ii) MOLECULAR TYPE: DNA(genomic)
- (iii) HYPOTHETICAL: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: *Chlamydomonas reinhardtii*
 - (B) STRAIN: RS-3
- (ix) FEATURE:
 - (A) NAME/KEY: -
 - (B) LOCATION: 1..3381
 - (C) OTHER INFORMATION: /note="encodes protoporphyrinogen oxidase"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

| | | | | | | |
|-------------|-------------|------------|------------|-------------|------------|-----|
| CTCGAGAGCG | TTGGAGGAAA | TCCGTTGGC | ACCTGTTCCG | GCTTCTTTGT | GTGCACGGCC | 60 |
| ACGTCCCCCT | TTCCTGCTAC | CCGCTCCCCC | CCGGCTTTAC | TGCCCTTCC | ACTCCTCGGC | 120 |
| TCCATCCCGA | TTCCATCCGC | TCCTCCTCCC | CCACCTAGAC | TGTCTACCGT | CTACCAGTTT | 180 |
| CTTGGGCAAT | CATTAACGTA | ACCCCGCCTC | CCTGCGCCTG | CCCCTCCCTC | CCTCTCCCCC | 240 |
| CCGCACAGCC | CGCCGCGGCC | GAGGCCCTGG | GCTCCTTCGA | CTACCCGCCG | ATGGGCGCCG | 300 |
| TGACGCTGTC | GTACCCGCTG | AGCGCCGTGC | GGGAGGAGCG | CAAGGCCTCG | GACGGGTCCG | 360 |
| TGCCGGGCTT | CGGTCAAGCTG | CACCCGCGCA | CGCAGGTGGG | CAAGTGCAGCG | CGTGTGCGG | 420 |
| GCAGGTGTGTT | GCAGGAGGGGA | GGGTGGTGGG | GGTTGGGGGT | GGGGGTGGGG | GGGATTGGGG | 480 |

| | | | | | | |
|------------|------------|-------------|------------|------------|------------|------|
| CGCTGGGTCG | TATCCCGCGG | TTGTATCCTC | GCGCTCCCCT | CATCCATTCC | CCCCTTCAAC | 540 |
| AACACACACG | GGCGCACACG | CACCCCTTT | GCGCTTACTT | TGTCTGGTGC | TCCTTAACAC | 600 |
| ACTCTTCGCT | TCATTTGGT | GTCTTCTAAC | ACACACACTT | GTCCACACAC | AGGGCATCAC | 660 |
| CACTCTGGC | ACCATCTACA | GCTCCAGCCT | GTTCCCCGGC | CGCGCGCCCG | AGGGCCACAT | 720 |
| GCTGCTGCTC | AACTACATCG | GCGGCACCAC | CAACCGCGGC | ATCGTCAACC | AGACCACCGA | 780 |
| GCAGCTGGTG | GAGCAGGTGT | GTGTGTGGGG | GGGTGGGGGG | GGGGCAGTGG | ATTTTTGGC | 840 |
| TGAGCCCCCT | GAGCAAAGCG | ATCCAGGGGG | GGCGAAGCCC | CCCAGGATTG | CCCCTGTCCG | 900 |
| TGCGTGCCTG | TGTGCCTGTG | TCGACAAAAAA | GTACCGTACT | GGCACAAACC | GCGAGTGC | 960 |
| CGTATTATTA | ATTGCAATT | CCTATTGTAG | AAAAATAGAC | GGCAGGGAAA | ACTCGGCCGG | 1020 |
| AGCGAGAACG | GACCTCGTGA | GTCCATGGAC | ATCTTGACTT | TCTTCAGTTC | GCGAGTATAG | 1080 |
| CTCTCGGCC | CTAAATATCT | TACATCCATG | TATCAAAACA | TGTCGACGAC | AAGCGTCTTG | 1140 |
| GGGCAAGAAT | GTCGAAATTG | TTTGCAACAG | CCAAACCATG | CGTCCCCGAG | CCTTACATGT | 1200 |
| GTCGCGGCC | GGGATCCCGC | GCCCGAGCCC | GGCTAGCCCT | TTGCGGTGCT | TGAGTGGGAT | 1260 |
| GTGGGTGAGG | TGCATTGGG | ATATCATGGA | CCGTGAAGTG | GCGTGGGTAA | GGTGGCGTGG | 1320 |
| CGTGGCGGGG | ACAGGGCATG | TCGGTGCCTC | GGCACAGCGT | TGGCCTAGTG | GCCAGTCCCG | 1380 |
| CTGGATGGC | TTGCAAGGGT | GCTGTTCATG | TCGCCGGTGC | CCATCGTCAC | ATCCCCTTGC | 1440 |
| GCTACATGGG | GCTCAGCCC | TTTCCAGCT | GTACAAAGCT | GACACCCCTT | GTTGTGTGGC | 1500 |
| GTCTTGGACC | CGTGTGCTT | CGGAGCTGGC | CAGAACCCCC | TGTGGGCACA | CACACGCACA | 1560 |
| CACACACACA | CACACACACA | CACACACACA | CACACACACA | CACACACACA | CACACACACA | 1620 |
| CACACACACA | CACACACACA | CACACACACA | CACACACACA | CACATTTCG | TCCTGCAGCC | 1680 |
| CCGAACCCCG | CCGCCCGTTC | CACGTCTTCC | ACCTGCCGCA | CCCCCCCCCC | TGCCGCACGC | 1740 |
| CTGCTCTCAC | CGCCTCTCCC | CCCACCCAT | CTCCCTGCAG | GTGGACAAGG | ACCTGCGCAA | 1800 |
| CATGGTCATC | AAGCCCGACG | CGCCCAAGCC | CCGTGTGGTG | GGCGTGC | TGTGGCCGCG | 1860 |
| CGCCATCCCG | CAGGTGTGAG | GGCGCAGCAG | CCGGAGGGAT | GGGCTAGATC | CTAGTTCTC | 1920 |
| AAAGAGCTCT | ACAGCCCTAT | AACCTCGACC | TGCGACCTTC | GACCTGATAA | CCTGGCTGCC | 1980 |
| CCCTCCCAAC | CTAGCCACCT | CTCCCCGGAT | TTGGGTTCAC | TCGGTTGACT | TGCTTTGGG | 2040 |
| TTCTGGAATC | AACTTCACCT | GTTGTATACT | TTGCTGC | ACTCTGTACC | ACTCTTGCA | 2100 |
| TTAGGTTCGG | TTTAGTTGG | GCTGCATGTG | TAACCCCTCC | TCCCCGCCCT | GCCACCTGCA | 2160 |

| | | | | | | |
|-------------|------------|------------|------------|------------|------------|------|
| GTTCAACCTG | GGCCACCTGG | AGCAGCTGGA | CAAGGCGCGC | AAGGCGCTGG | ACGCGGCCGG | 2220 |
| GCTGCAGGGC | GTGCACCTGG | GGGGCAACTA | CGTCAGCGGT | GAGCGCGTGG | GCAGCAGCAG | 2280 |
| CAGCAGGAAG | AGGGGAGGGG | AGGGGAGGGG | AGGGTACAAG | GAGGAGGTTG | AGCAGGAGGT | 2340 |
| GGTGCTAAGG | CGCAAAGCAA | GGCGGTGTTG | TATCCTCATT | GAUTGAAACC | GGGAAACCCA | 2400 |
| GCATGAACAA | GAGGTCAAGG | GAUTGCAAGG | AGCGGAGGCT | ACATGTATGA | CTACCCCCGA | 2460 |
| CGCGGGCGAT | GATTCCCTGA | CTATTGGGAC | CTATTCGTT | GGGCTCGGGC | ACATGACCCC | 2520 |
| CCTGGCCCT | TCGCTGTATG | GTGCCCAGCC | GCCCAGCCGC | CCCCCGCCCA | CACGTGTGCC | 2580 |
| CACGCCTTG | CCTCATCCCC | AACCCCTCG | GCCCCTCTCC | CCCCTCGAAC | CCCTGCAACC | 2640 |
| AGGTGTGGCC | CTGGGCAAGG | TGGTGGAGCA | CGGCTACGAG | TCCGCAGCCA | ACCTGGCCAA | 2700 |
| GAGCGTGTCC | AAGGCCGAG | TCAAGGCCTA | AGCGGCTGCA | GCAGTAGCAG | CAGCAGCATC | 2760 |
| GGGCTGTAGC | TGGTAAATGC | CGCAGTGGCA | CCGGCAGCAG | CAATTGGCAA | GCACTTGGGG | 2820 |
| CAAGCGGAGT | GGAGGCGAGG | GGGGGGCTAC | CATTGGCGCT | TGCTGGGATG | TGTAGTAACA | 2880 |
| GTTGGAATGG | ATCGGGGATG | TGGAGCTAGG | GGTCGGGGG | TCTGCCAAGG | ACATAGGTGG | 2940 |
| TGCTGGGATG | AGCGATGTGG | TTGGTAAAGC | TCTGTCGGCA | CCGTTATGTG | CGGGTTAACT | 3000 |
| GCACTATGAC | GCTCCGTTGT | ACAGCCCCGT | TGTGCATTGT | TTGCATGAAG | TTTGGCGAG | 3060 |
| AGTGAGTTGG | CGCACACGCG | GGCGGTTTG | GGGGCACTGT | CCCTCAGTGT | GGTCCCAGCA | 3120 |
| TAGCACAGGA | GAGACACAGA | ACTGAGTGAC | ATAGACTAGG | TCTCGAAGTA | CCTTCAAAAG | 3180 |
| GGGGCTATAA | ATTGCGAATA | CCCGGAGCAG | GGGCCAGAC | CCAAGGCATT | GAUTGTCAGT | 3240 |
| GCACAAGCGA | AAGACCAATT | GCATGGGTTG | CTTCCGTGGT | GGGAAGAGGA | GGGCAGGGGA | 3300 |
| GCATCGTCAG | GTGTATGTTG | CGGCTTCGCC | CATAAGTGCC | ATGGTTTCGA | AGATGCTTAA | 3360 |
| GAUTAACAAAT | GCCAACTCGA | G | | | | 3381 |

(2) INFORMATION FOR SEQ ID NO:20

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 18 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULAR TYPE: oligonucleotide
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO

(ix) FEATURE:

- (A) NAME/KEY: -
- (B) LOCATION: 1..18

(C) OTHER INFORMATION: /note="oligonucleotide primer 3A for *Chlamydomonas reinhardtii*"

(xi) SEQUENCE DESCRIPTION:SEQ ID NO:20

TTCCACAGTCT TCCACCTG

(2) INFORMATION FOR SEQ ID NO:21

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: oligonucleotide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(ix) FEATURE:

- (A) NAME/KEY: -
- (B) LOCATION: 1..17

(C) OTHER INFORMATION: /note="oligonucleotide primer 3B for *Chlamydomonas reinhardtii*"

(xi) SEQUENCE DESCRIPTION:SEQ ID NO:21

CTAGGATCTA GCCCATC

(2) INFORMATION FOR SEQ ID NO:22

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: oligonucleotide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

- (A) NAME/KEY: -
- (B) LOCATION: 1..18

(C) OTHER INFORMATION: /note="oligonucleotide primer 4A for *Chlamydomonas reinhardtii*"

(xi) SEQUENCE DESCRIPTION:SEQ ID NO:22

CTGCATGTGT AACCCCTC

(2) INFORMATION FOR SEQ ID NO:23

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: oligonucleotide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(ix) FEATURE:

- (A) NAME/KEY: -
- (B) LOCATION: 1..18

(C) OTHER INFORMATION: /note="oligonucleotide primer 4B for *Chlamydomonas reinhardtii*"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23

GACCTCTTGT TCATGCTG

(2) INFORMATION FOR SEQ ID NO:24

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: oligonucleotide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(ix) FEATURE:

- (A) NAME/KEY: -
- (B) LOCATION: 1..17

(C) OTHER INFORMATION: /note="oligonucleotide primer 5B for *Chlamydomonas reinhardtii*"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24

CGGCATTTAC CAGCTAC